Acoustic aspects of vowel harmony in French [Congress of Phonetic Sciences (ICPhS)]

Noël Nguyen, Zsuzsanna Fagyal

Acoustic aspects of vowel harmony in French

Noël Nguyen* and **Zsuzsanna Fagyal†**

* Laboratoire Parole & Langage, CNRS & Université de Provence, France

† Department of French, University of Illinois at Urbana-Champaign, USA

nguyen@lpl.univ-aix.fr, zsfagyal@uiuc.edu

ABSTRACT

This paper explores acoustic and articulatory aspects of regressive vowel-to-vowel assimilation known as vowel harmony (VH) in French. Based on three speakers' renditions of 136 pairs of disyllabic word pairs containing a mid-vowel in the first, and a low or a non-low vowel in the second, syllables of each pair, we examined assimilatory effects of final vowels on the duration and spectral properties of non-final mid-vowels. Results show that /e/ and /o/ have longer duration, and occupy a more peripheral position in two speakers' vowel spaces when followed by a non-low rather than a low vowel. Phonological implications of these findings could be that vocalic contrasts referred to as tense-lax distinction in other languages and varieties of French would allow characterizing the assimilatory behavior of mid vowels of standard French in a uniform way.

1 INTRODUCTION

In French, the quality of mid vowels is known to vary to a greater extent in non-final syllables than in final syllables. This variability has been ascribed to several factors, which include the syllable structure, the identity of the post-vocalic consonant, the morpho-phonological structure of the word, and vowel harmony. Vowel harmony in French is described as a word-level anticipatory process affecting non-final mid vowels that assimilate in height to the final tonic vowel. The non-final vowel will tend to be mid-high before a high or mid-high vowel (e.g. *aimer* [eme] "to love"), and mid-low before a low or mid-low vowel (*aimable* [ɛmabl] "kind", see [1]).

Most authors consider that vowel harmony (henceforth, VH) is optional, possibly speaker-dependent, and that it is more likely to occur in conversational speech than in formal speech [2, 3, 1]. VH is also often said to be restricted to the front unrounded pair /e-ɛ/ [2, 3, 4], although Grammont [5] extends it to /œ/ (see also [6]), and Tranel [7] to both /œ/ and /ɔ/ when these are followed by their mid-high counterparts /ø/ and /o/. In addition, it is frequently argued that VH mainly affects mid-low vowels followed by a high or mid-high vowel [6, 3, 5, 4, 7]. A noticeable exception is Dell [2], who suggests that /e/ can be realized as [ɛ] before a low vowel.

Vowel harmony is sometimes portrayed as being sensitive to morphological factors, although there seems to be no general agreement on this point. Dell's [2] HARM rule states that the assimilating and the assimilated vowels have to belong to separate morphemes. On this account, VH would never happen in monomorphemic words (hence, *récolte* "harvest" can only be realized as [ʁekɔlt], as opposed to *[ʁɛkɔlt]). Tranel [7], conversely, suggests that VH applies regardless of the presence or absence of a morpheme boundary between the two adjacent vowels (e.g. *auto*, 1 morph., "car", realized as [oto]). Note, however, that most examples of VH given in the literature involve bimorphemic words, with a vowel in a derivational or inflectional suffix exerting a regressive influence on the vowel in the root.

Previous studies on VH have mostly focused on the effect of the word-final vowel on the penultimate vowel, and little attention has been paid to whether VH can extend over a larger domain in polysyllabic words, although Grammont [5] implicitly suggests that this might indeed be the case (e.g. *bégaiement* [bɛgɛmɑ̃] "stuttering" vs. *bégayer* [begeje] "to stutter"). There is also little information available on the potential effect of regional accent on patterns of VH, as most authors confine the scope of their study to Parisian French.

To our knowledge, and although VH is mentioned in most textbooks on the phonetics and phonology of French, the articulatory and acoustic aspects of VH have not yet been systematically explored. The present study aimed at identifying the acoustic correlates of VH in French. Our goal was to characterize the effect of VH on the spectral shape of a non-final mid vowel. We also sought to determine whether there are differences between front mid vowels and back mid vowels, with regards to their sensitivity to VH.

2 METHOD

The corpus consisted of 136 pairs of disyllabic nouns, adjectives, and infinitives. The first syllable always contained a mid vowel (henceworth, V1) and was phonemically identical in both words of the pair. The second syllable contained a non-low vowel in one word and a low vowel in the other word of the pair (hence-

worth, V2 for both). The syllable onset was either a single consonant or a consonant cluster. We restrict our analysis to words in which the first syllable was open (119 pairs). Table 1 shows the sets of word pairs with one example for each set, the phonemic identity of V1, the phonemic identity of V2, and the total number of word pairs in each set.

word-pair set	V1	V2	n
été /ete/ - *éther* /etɛʁ/	e	e/ɛ	18
prêteuse /pʁetøz/ - *prêteur* /pʁetœʁ/	e	ø/œ	8
dévot /devo/ - *dévote* /devɔt/	e	o/ɔ	6
potée /pote/ - *poterne* /potɛʁn/	o	e/ɛ	34
poseuse /pozøz/ - *poseur* /pozœʁ/	o	ø/œ	19
auto /oto/ - *automne* /otɔn/	o	o/ɔ	4
épice /epis/ - *épate* /epat/	e	i/a	14
notice /notis/ - *nota* /nota/	o	i/a	14

Table 1: Types and number of word pairs.

V1s are systematically transcribed in Table 1 as high-mid vowels. This is consistent with the *loi de position*, which we somewhat arbitrarily adopted as a rule of transcription. The actual quality of V1 was of course an empirical issue that this acoustic study was precisely intended to address. There was a morphological (derivational) relationship between the two members of the pair for part of the word pairs, e.g. *prêteuse-prêteur*, whereas other word pairs were morphologically unrelated. The potential influence of morphology on VH is not examined in this paper.

The corpus[1] was recorded by six speakers representing two dialects of French. Results are presented here for three of these speakers, referred to as S1 (female, aged 28 y.), S2 (female, aged 34 y.) and S3 (male, aged 39 y.). S1 is a native speaker of the Northern variety of French spoken in Paris, whereas S2 and S3 represent the Southern variety of French spoken in the area of Aix-en-Provence. Note, however, that S2 and S3 do not have post-consonantal final schwa, and just like in standard French, they pronounce mid-high vowels in final closed syllables ending in /z/ (e.g. *prêteuse*, see Table 1).

The recordings took place in the Phonetics Laboratory of UIUC for S1 and at the LPL, Aix-en-Provence, for S2 and S3. At both locations, the recordings were carried out in an anechoic chamber using a high-quality microphone and a DAT recorder. The material was presented to the speaker on slides shown on a computer screen. Each word appeared twice on each slide, in a carrier sentence as in the following example: "Il retape *été* parfois; *été*" /il ʁətap ete paʁfwa ‖ ete/ "he retypes *summer* sometimes; *summer*". The utterance-final target words were intended as stimuli in a companion perceptual experiment, but are also incorporated in this acoustic study. The speakers were asked to read each sentence at a normal rate, with no special emphasis on any word, and no pause between words in the first phrase (up to *parfois*). The entire set of sentences and the isolated words were recorded in a random order twice, in two separate sessions.

The acoustic data were manually segmented on the basis of the waveform and of the corresponding wide-band spectrogram. For each target word, markers were placed at the onset and offset of the first and second vowels. The frequencies of the first three formants were automatically measured every 5 ms over the entire utterance using the ESPS function *formant*. The frequencies of F1 and F2 were extracted at the acoustic midpoint of the first vowel and of the second vowel. These measures were then verified manually. For a given speaker and a given combination of vowels, extreme formant frequency values were checked and corrected (when this proved necessary) using both an FFT spectrum and an LPC spectrum computed over a 50-ms window centered at the midpoint of the vowel, together with a wide-band spectrogram.

Pairwise comparisons were made between the words containing a final non-low vowel (e.g. *été*), on the one hand, and those with a final low vowel(*éther*), on the other hand. Since our survey of the literature showed that vowel harmony generally means regressive assimilation in height, and given that F1 is inversely related to vowel height, we predicted that F1 frequency should be lower in V1 when V2 is mid-high or high as opposed to mid-low or low.

3 RESULTS

Table 2 shows mean differences in formant frequencies for V1 depending on whether V2 is non-low or low, for each V1 category and each speaker. The difference is positive when the formant frequency is higher before a non-low vowel than before a low vowel. For example, F2 was on average 61 Hz higher for /e/ followed by /e/ (e.g. *été*) as opposed to /ɛ/ (*éther*) for Speaker S1. For the sake of clarity, only statistically significant differences ($p < .05$, in two-tailed, paired t-tests) are shown.

Table 2 reveals that, in keeping with our expectations, F1 frequency tended to be lower before a non-low vowel than before a low vowel, although this trend appeared to be specific to a subset of V1-V2 combinations. In general, F2 frequency was higher for /e/ prior to a non-low vowel compared to a low vowel, and for /o/ before /i/ as opposed to /a/. In words containing /o/ followed by a mid vowel, F2 was generally lower in frequency when V2 was mid-high rather than mid-low. Differences between speakers were also observed, with Speaker S2 showing less systematic V2-dependent spectral variations in V1 than the two other speakers.

The variations in formant frequency for the first vowel,

[1] We thank P. Boula de Mareüil for his contribution to preparing the corpus.

Sp.	V1	Fmt	V2			
			e/ɛ	ø/œ	o/ɔ	i/a
S1	e	ΔF1	–14			–32
		ΔF2	+61	+85	+72	+117
	o	ΔF1				
		ΔF2	–55	–53	–65	+39
S2	e	ΔF1				–12
		ΔF2				+66
	o	ΔF1		–10		
		ΔF2	–29		–102	+30
S3	e	ΔF1	–22		–9	–32
		ΔF2	+43			+96
	o	ΔF1	–28	–10	–40	–40
		ΔF2	–43		–98	+46

Table 2: Mean difference in formant frequency (Hz) for V1 depending on V2 for each speaker. See text for details.

depending on the second vowel, are depicted in Figure 1 for Speaker S1. Each panel from (a) to (d) contains two pairs of 1-sigma ellipses associated with the two categories of V1, /e/ (in the high F2-frequency range) and /o/ (in the low F2-frequency range) respectively. For each category of V1, grey ellipses reflect the dispersion of F1 and F2 values before a non-low V2, and white ellipses show the dispersion of F1 and F2 before a low V2. Each panel corresponds to a specific V2 pair: /e-ɛ/ for Panel (a), /o-ɔ/ for Panel (b), /ø-œ/ for Panel (c), /i-a/ for Panel (d). Panel (e) shows the mean F1 and F2 frequencies for /i,e,ɛ,a,ɔ,o/ in word-final position. This vowel diagram was constructed to facilitate the interpretation of the four other panels, and also appears as a reference for these panels.

Figure 1 indicates that, in non-final syllables, both /e/ and /o/ generally tend to have a more peripheral position in the vowel space (i.e. to have more extreme F2 and, in some cases, F1 values) prior to a non-low vowel compared to a low vowel. An exception to this tendency is /o/ combined with the /i-a/ pair (Panel (d)).

Sp.	V1	V2			
		e/ɛ	ø/œ	o/ɔ	i/a
S1	e	+4		+8	+5
	o	+5	+5	+7	
S2	e	+11			+6
	o	+11	+6	+10	+8
S3	e	+13	+6	+10	+5
	o	+10		+16	

Table 3: Mean difference in the duration (ms) of V1 depending on V2 for each speaker. See text for details.

We also measured the duration of the non-final mid vowel in each word, with a view to determine whether this duration would vary contingent upon the word-final vowel. Table 3 shows the mean difference in the duration of V1 depending on whether V2 is non-low

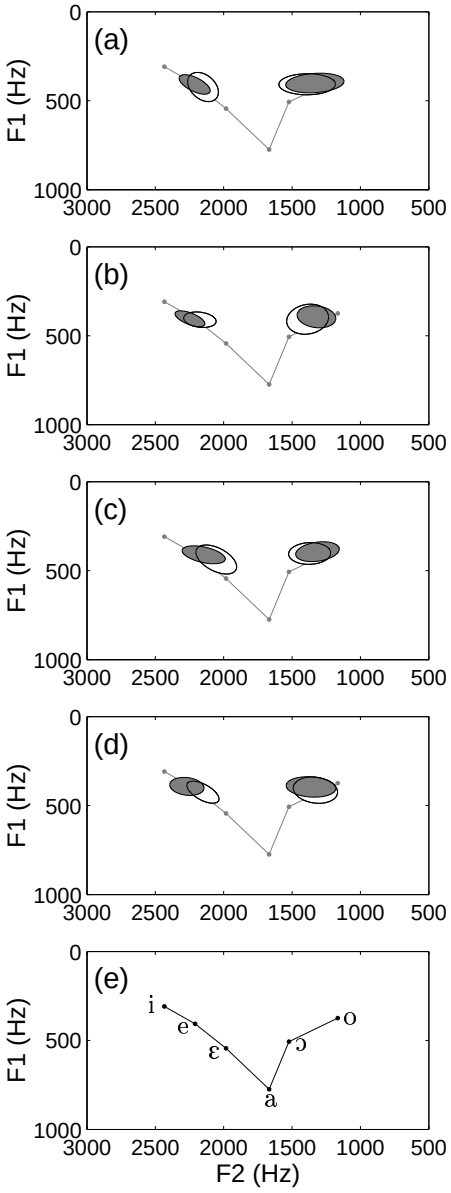

Figure 1: Variations in F1 and F2 frequency for V1 depending on V2 for Speaker S1. See text for details.

or low, for each V1 category and each speaker. This difference is positive when V1 is longer before a non-low vowel than before a low vowel. Only statistically significant differences ($p < .05$) are reported. Table 3 demonstrates that for a majority of V1-V2 combinations, V1 had a longer duration when V2 was non-low rather than low. These durational differences were small but they were highly significant in most cases.

4 DISCUSSION

In summary, both spectral and durational differences were found in mid vowels depending on the following, word-final vowel. Mid vowels tended to be slightly longer, and have a lower F1 frequency and/or a more extreme F2 frequency before a mid-high vowel than

before a mid-low vowel. Such a tendency was more marked for two speakers out of three. Contrary to the assumption generally made in the literature that vowel harmony only applies to the /e-ɛ/ pair, V2-dependent acoustic variations were observed for both front and back mid vowels.

One important issue is whether our data are consistent with the notion that vowel harmony in French is an assimilation in vowel height, as has been assumed in previous work. The fact that, for Speakers S1 and S2, F1 frequency in the first vowel was less sensitive than F2 frequency to the quality of the following vowel does not seem at a first glance to support this assumption. However, Figure 1(e) shows that what we refer to as vowel height is associated with variations in both F1 and F2 frequency. Thus, moving from /e/ to /ɛ/ involves a downward shift in F2 along with an increase in F1. Reciprocally, moving from /o/ to /ɔ/ results in a substantial upward shift in F2 in conjunction with a raising of F1. In that respect, our data fit with an account of VH in terms of assimilation in vowel height. An alternative approach, however, would be to interpret the observed VH patterns as involving the tense/lax distinction. As is well-known, tense vowels such as /e,o/ have more extreme displacements of the formants from their "neutral" values than the corresponding lax vowels (/ɛ,ɔ/) [8]. If we assume that the tenseness/laxness of a word-final vowel extends to the preceding vowel, we would expect the latter to reside closer to the periphery of the vowel space when the word-final vowel is tense as opposed to lax. This was indeed the general tendency. Thus, the tense/lax distinction may allow the behavior of both the front and back mid vowels in VH to be characterized in a uniform way. In addition, this distinction is associated with a difference in duration, with tense vowels being generally longer than their lax counterparts, and that may account for why V1 varied in duration depending on V2 in this study (Table 3).

Besides empirical evidence, i.e. actual spectral and durational measures, our proposal to interpret the assimilatory behavior of mid vowels in our data in terms of a tense/lax distinction is supported by several phonological analyses of the Standard French vowel system. [9], for example, raises the possibility of a tense/lax distinction with respect to rimes in Standard French, while studies such as [10] review the hypothesis that historical developments of the modern vowel system of Standard French involved moving from a length-based distinction to a "length and quality"-based distinction.

Although our spectral analyses were performed at the acoustic mid-point of each vowel only, informal observations suggest that VH effects may also encompass dynamic changes in the shape of the spectrum during the vowel. To give an example, spectrograms of *coder* [kode] "to code" and *codex* [kodɛks] "pharmacopoiea" (first 400 ms for both words) are shown in Figure 2. Differences both in the duration of V1 and in the trajectory of F2 and F3 in V1 can be noted. In *coder*, F2 has a flat trajectory over the first part of the vowel, and increases in frequency after the vowel's mid-point. In *codex*, F2 steadily raises throughout the vowel. These variations are partially captured by the spectra associated with V1 for both words, with F2 frequency being lower for *coder* than for *codex*. However, spectral variations over time conditioned by vowel harmony remain to be investigated in a systematic manner.

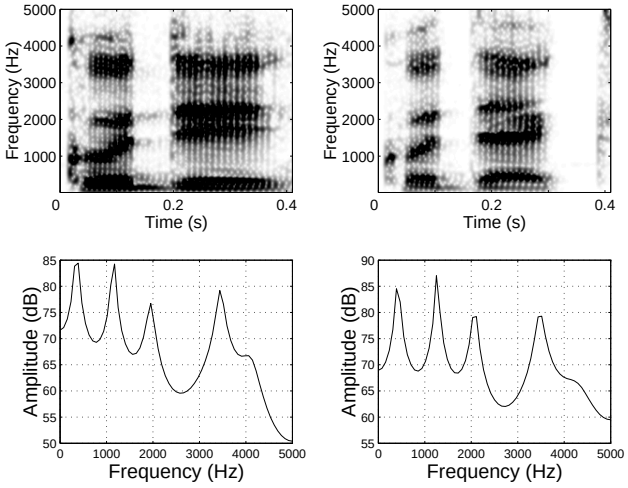

Figure 2: Top: spectrograms of the words *coder* (left) and *codex* (right) spoken by S3. Bottom: Spectral shape of V1 in each of the two words.

REFERENCES

[1] D.C. Walker, *French sound structure*, Univ. Calgary Press, Calgary, 2001.

[2] F. Dell, *Les règles et les sons*, Hermann, Paris, 1972.

[3] P. Fouché, *Traité de prononciation française*, Klincksieck, Paris, 1956.

[4] B. Malmberg, *Phonétique française*, Hermods, Malmö, 1969.

[5] M. Grammont, *Traité de phonétique*, Delagrave, Paris, 1933.

[6] J. Casagrande, *The sound system of French*, Georgetown Univ. Press, Washington, D.C., 1984.

[7] B. Tranel, *The Sounds of French: an introduction*, Cambridge Univ. Press, Cambridge, UK, 1987.

[8] K.N. Stevens, *Acoustic Phonetics*, MIT Press, Cambridge, Mass., 1998.

[9] M. Plénat, "On the structure of rime in Standard French," *Linguistics*, vol. 25, pp. 867–887, 1987.

[10] Y.-C. Morin, M.-C. Langlois, and Varin M.-E., "Tensing of word-final [c] to [o] in French: The phonologization of a morpho-phonological rule," *Romance Philology*, vol. XLIII, pp. 507–528, 1990.